パーセプトロンのノイズ有り学習(基研研究会「ニューラルネットワーク〜これからの統計力学的アプローチ〜」,研究会報告)

パーセプトロンのノイズ有り学習

奈良女子大学 理学部 物理科学科 上江洌 達也

1. Introduction

現実の系には必ず何らかのノイズが存在するため、学習の問題においても、ノイズの効果を調べることは極めて意味深い。ここでは、教師パーセプトロンに外部ノイズが加わった時の生徒パーセプトロンによる学習の問題を、統計力学的手法を用いて解析する。

2. 一層パーセプトロンの学習

2.1. 問題設定

教師パーセプトロンの荷重ベクトル $\boldsymbol{w}^o$、生徒パーセプトロンの荷重ベクトル $\boldsymbol{w}$ 、例題ベクトル $\boldsymbol{x}$ を いずれも、N 次元ベクトルとし、その大きさは $\sqrt{N}$ であるとする。μ番目の例題 $\boldsymbol{x}^\mu$に対する生徒の出力 r^μ は、

$$r^\mu = \mathrm{sgn}(h^\mu),\ \ h^\mu \equiv (\boldsymbol{x}^\mu \cdot \boldsymbol{w})/\sqrt{N}$$

である。一方、 μ番目の例題 $\boldsymbol{x}^\mu$に対して、教師が $r^{o\mu} = +1$ を出力する確率を次の形に仮定する。

$$p_r(+1|\boldsymbol{x}^\mu) \ = \ \mathcal{P}(h^{o\mu}) = \frac{1 + P(h^{o\mu})}{2},\ \ h^{o\mu} \equiv (\boldsymbol{x}^\mu \cdot \boldsymbol{w}^o)/\sqrt{N}.$$

学習アルゴリズムとしては、ここでは、Gibbs アルゴリズムを用いる。p 個の例題とそれに対する教師出力の組、training set

$$\{\boldsymbol{x}^\mu, R^\mu\}, \mu = 1, \cdots, p,$$

が与えられたとき、生徒 $\boldsymbol{w}$ のエネルギー $E_p[w, x]$ を、p 個の出力のうち生徒出力と教師出力が異なる回数と定義する。そして、生徒を $\exp(-\beta E_p[w, x])$ に比例する確率で選択する。$\beta = 1/T$ で、T は、「温度」。特に、$T \to +0$ は、誤り最小アルゴリズム (Minimum error algorithm) となる。

学習の到達度を測るために、汎化誤差 ϵ_g を計算する。これは、新しい例題が与えられたときに、教師出力と異なる出力を出す確率の平均値である。

$$\epsilon_g = < \mathcal{P}(u^o)(1 - \Theta(u)) + (1 - \mathcal{P}(u^o))\Theta(u) >_x .$$

汎化誤差の振舞いは、荷重ベクトルの種類や、ノイズの入り方等に依存している。ここでは、次のような場合を取り扱う。

○荷重ベクトルの種類。

・連続的な場合 (Spherical case)。各成分が連続変数。・離散的な場合 (Ising case)。$w_i = \pm 1$.

○ノイズのタイプ

・入力ノイズ。 入力される例題ベクトルにノイズが入る。平均 0 、分散 σ のノイズ η_iが、各成分 x_i に独立にはいるとすると、例題 $\boldsymbol{x}$ にたいして、$r^o = +1$ となる確率 $p_r(+1|\boldsymbol{x})$ は、

$$p_r(+1|\boldsymbol{x}) \ = \ H[-(\boldsymbol{x} \cdot \boldsymbol{w}^o)/(\sigma\sqrt{N})],\ \ H(u) = \int_u^\infty dx \exp[-x^2/2]/\sqrt{2\pi}.$$

・出力ノイズ。 教師出力が、確率 λ で、反転する。

$$p_r(+1|\boldsymbol{x}) = \lambda + (1 - 2\lambda)\Theta[(\boldsymbol{x} \cdot \boldsymbol{w}^0)/\sqrt{N}]$$

・一般的に、次のようにかける。

$$p_r(+1|\boldsymbol{x}) \;=\; \mathcal{P}(h^o) = \frac{1+P(h^o)}{2}, \;\; h^o \equiv (\boldsymbol{x}\cdot\boldsymbol{w}^o)/\sqrt{N}.$$

$P(u)$ の $u=0$ 近傍での振舞いを $P(u) \simeq a\,\mathrm{sgn}(u)|u|^\delta$, とし、このクラスを $P_\delta(u)$ と表す。
入力ノイズは、$\delta=1$ であり、 出力ノイズ は $\delta=0$ である。
　このとき、汎化誤差は、

$$\begin{aligned}\epsilon_g &= \epsilon_{min} + \int_0^\infty DyH(y/\zeta)[P(y)-P(-y)],\\ \epsilon_{min} &= \frac{1}{2}-\frac{1}{2}\int_0^\infty Dy[P(y)-P(-y)] = \int_0^\infty Dy(1-\mathcal{P}(y)) + \int_{-\infty}^0 Dy\mathcal{P}(y),\end{aligned}$$

となる。ここで、 $Dy=\exp(-y^2/2)/\sqrt{2\pi}$, $\zeta=\frac{\sqrt{1-R^2}}{R}$. ϵ_{min} は、$\boldsymbol{w}=\boldsymbol{w}^o$ のときの、汎化誤差。R は、教師と生徒のオーバーラップ。ϵ_g は R の単調減少関数となる。

2.2. R の計算

レプリカ法を用いて R を計算する。我々は、RS 解、即ち、全ての量が、レプリカ指数 α,β に依存しない解と、この解が不安定な場合は、1RSB 解, 即ち、レプリカ対称性を一回破る解 を考える。具体的計算は、文献 [1, 2, 3] を参照。

2.3. 結果

　鞍点方程式を解いて、R を計算することにより、汎化誤差が求まる。結果は、ノイズのタイプや、荷重ベクトルの種類によって、2つに分類される。

$$\begin{aligned}\Delta\epsilon_g &\equiv \epsilon_g-\epsilon_{min} \propto \alpha^{-\gamma}, \quad \alpha \gg 1 \text{ のとき。}\\ \Delta\epsilon_g &= 0\,(\text{完全学習}), \quad \alpha > \alpha_p \text{ のとき。}\end{aligned}$$

I. 一層、$\boldsymbol{w}^o, \boldsymbol{w}$ 共に連続、関数族 P_δ
汎化誤差は、常に、ベキ的。

$$\Delta\epsilon_g = \alpha^{-\gamma}.$$

1. ギブスアルゴリズム及び 誤り最少アルゴリズム。
   $\Delta\epsilon_g$ の漸近形

	RS 解	1RSB 解
出力ノイズ	α^{-1}	α^{-1}
入力ノイズ	$\alpha^{-\frac{1}{2}}$	$\alpha^{-\frac{2}{3}}(\ln\alpha)^{\frac{1}{3}}$
P_δ $(\delta>0)$	$\alpha^{-\frac{1+\delta}{1+3\delta}}$	$\alpha^{-\frac{1+\delta}{1+2\delta}}(\ln\alpha)^{\frac{1+\delta}{2(1+2\delta)}}$

2. 最適温度の存在について。
   ・出力ノイズの場合

$$\Delta\epsilon_g \;\simeq\; \psi_0^{(RS)}(T)\alpha^{-1}: \;\text{RS 解}, \;\; \Delta\epsilon_g \simeq \psi_0^{(1RSB)}(T)\alpha^{-1}: \;\text{1RSB 解}.$$

$\psi_0^{(RS)}(T)$ は、T_0^* で最小値となり、このとき、RS 解は安定。1RSB 解を考慮しても最小である。
・入力ノイズの場合

$$\begin{aligned}\Delta\epsilon_g &\simeq \{\psi_1^{(RS)}(T)\}^{\frac{2(1+\delta)}{1+3\delta}}\alpha^{-\frac{1+\delta}{1+3\delta}}: \;\text{RS 解},\\ \Delta\epsilon_g &\simeq \psi_\delta^{(1RSB)}(T)\alpha^{-\frac{1+\delta}{1+2\delta}}(\ln\alpha)^{\frac{1+\delta}{2(1+2\delta)}}: \;\text{1RSB 解}.\end{aligned}$$

RS 解は、

$$\psi_1^{(RS)}(T) \propto \sqrt{1/T} \ \text{for} \ T \ll \alpha^{-1},$$

となるが、この解は、漸近的に不安定。一方、1RSB 解は、

$$\psi_\delta^{(1RSB)}(T) \propto \delta^{\frac{1+\delta}{2(1+2\delta)}} \ \text{for} \ \alpha^{-\frac{\delta}{3(1+2\delta)}} \ll T \ll \alpha^{\frac{\delta}{1+2\delta}}$$

となり、 広い温度領域で一定。

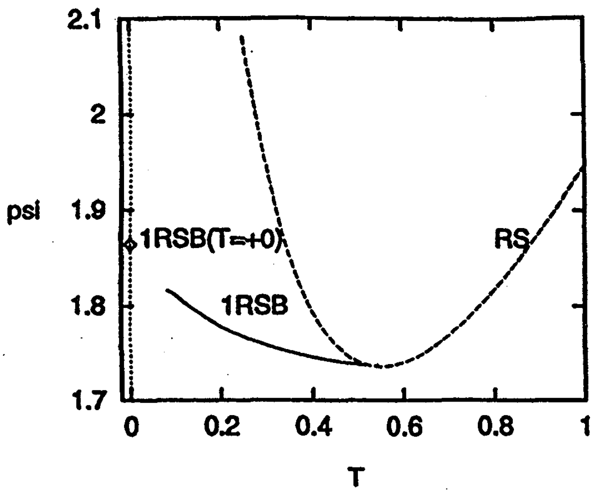

図 1: 出力ノイズの学習曲線のα^{-1}の係数の温度依存性。$P(u) = 0.5 \ \mathrm{sgn}\ (u)$。点線と実線は、各々$\psi_0^{(RS)}(T)$ と $\psi_0^{(1RSB)}(T)$。菱形は誤り最小アルゴリズムでの 1RSB 解の係数 $\psi_{MEA}^{(1RSB)}$ 。

II. 一層、 $\boldsymbol{w}^o, \boldsymbol{w}$ 共に離散的、関数族 P_δ、
$\Delta\epsilon_g$ の漸近形
RS 解は、漸近的に不安定。

	1RSB 解
$\delta > \frac{1}{2}$	$(\ln \alpha/\alpha)^{\frac{1+\delta}{2\delta-1}}$
$\delta = \frac{1}{2}$	$e^{-3F_0\alpha}$, F_0 は 定数
$\delta < \frac{1}{2}$	完全学習

3. 二層パーセプトロンの学習 – non-overlapping case –

3.1. 問題設定

入力層 $N = MK$個、中間層 K個、出力層1個のパーセプトロンを扱う[4]。教師パーセプトロンの荷重ベクトル $\boldsymbol{w}_l^t$ $(l = 1, \cdots K)$, 生徒パーセプトロンの荷重ベクトル $\boldsymbol{w}_l^s$ $(l = 1, \cdots K)$, 例題 $\boldsymbol{x}_l$ のいずれも、M 次元ベクトルで、大きさは $\sqrt{M}$とする。

μ番目の例題 $\{\boldsymbol{x}_l^\mu\}$ に対する教師出力 σ_t^μは、ノイズ無しのとき、

$$\sigma_t^\mu = B_t(\{\sigma_l^\mu\}), \quad \sigma_l^\mu = \mathrm{sgn}(h_{l,\mu}^t), \quad h_{l,\mu}^t \equiv (\boldsymbol{x}_l^\mu \cdot \boldsymbol{w}_l^t)/\sqrt{M}.$$

とする。ここで、$B_t(\{\sigma_l^\mu\})$ は、教師を特徴づける Boolean function。教師信号にノイズが入るときは、教師が σ_t^μ を出力する確率を次のように表す。

$$p_r(\sigma_t^\mu|\{\boldsymbol{x}_l^\mu\}) = \mathcal{P}(\sigma_t^\mu|\{h_{l,\mu}^t\}).$$

μ番目の例題 $\{\boldsymbol{x}_l^\mu\}$ に対する生徒の出力 σ_s^μは、

$$\sigma_s^\mu = B_s(\{\sigma_l'^\mu\}), \quad \sigma_l'^\mu = \mathrm{sgn}(h_{l,\mu}^s), \quad h_{l,\mu}^s \equiv (\boldsymbol{x}_l^\mu \cdot \boldsymbol{w}_l^s)/\sqrt{M}.$$

ここで、$B_s(\{\sigma_l'^\mu\})$ は、生徒の出力を特徴づける Boolean function。

3.2. モデルとその結果

レプリカ法を用いて、鞍点方程式を求め、R を計算することにより、汎化誤差が求まる。
$\boldsymbol{w}^t, \boldsymbol{w}^s$ は、共に連続とし、ノイズは、出力ノイズと入力ノイズを扱う。また、RS 解のみを考える。
更に、$B_t = B_s = B$ とし、permutation symmetry を仮定する。

結果

$\Delta\epsilon_g$ の漸近形

	RS 解
決定論的	α^{-1}
出力ノイズ	α^{-1}
入力ノイズ	$\alpha^{-1/2}$

参考文献

[1] T. Uezu and Y. Kabashima, J. of Phys. **A29**(1996),L55: T. Uezu, Y. Kabashima, K. Nokura and N. Nakamura, J. of Phys. Soc. Japan **65**(1996),3797.

[2] T. Uezu and Y. Kabashima, J. Phys. A: Math. Gen.**29**(1996),L439.

[3] T. Uezu, J. Phys. A: Math. Gen.**30**(1997),L777.

[4] B. Schottky, J. Phys. **A28**(1995),4515: B. Schottky and U. Krey, J. Phys. A: Math. Gen.**30**(1997),8541.